

395747

IN SEARCH OF RANDOM NOISE

DO KESTER

*Laboratory for Space Research**Postbus 800, 9700 AV Groningen, The Netherlands*TJ. ROMKE BONTEKOE¹*Astrophysics Division of ESA, ESTEC, Postbus 299, 2200 AG Noordwijk,
The Netherlands*

ABSTRACT In order to make the best high resolution images of IRAS data it is necessary to incorporate any knowledge about the instrument into a model: the IRAS model. This is necessary since every remaining systematic effect will be amplified by any high resolution technique into spurious artifacts in the images. The search for random noise is in fact the never-ending quest for better quality results, and can only be obtained by better models.

The Dutch high-resolution effort has resulted in HIRAS which drives the MEMSYS5² algorithm. It is specifically designed for IRAS image construction. A detailed description of HIRAS with many results is in preparation (Bontekoe, Koper & Kester 1993). In this paper we emphasize many of the instrumental effects incorporated in the IRAS model, including our improved 100 μm IRAS response functions.

INTRODUCTION

The InfraRed Astronomical Satellite (IRAS) All Sky Survey was designed and optimized for the detection of point sources. This allowed the survey to be conducted in the form of narrow strip scans with redundant coverage of the sky, but with non-uniform covering densities. The data show in addition to point sources many sources of extended emission, which are best analyzed from images. However, the non-uniform coverage now forms a significant obstacle in the image (re-)construction. Low resolution images, such as in the Infrared Sky Survey Atlas, yield spatial resolutions of five to ten times the IRAS telescope diffraction limit; HIRAS can improve on this by a factor of one to two times!

In HIRAS the imaging equation

$$d = R f \pm \sigma, \quad (1)$$

¹Current address: Bontekoe Data Consultancy, Jaap Bergmanstraat 3, 2221 BM Katwijk ZH, The Netherlands

²MemSys5 is a software package of Maximum Entropy Data Consultants Ltd. Cambridge, UK.

is solved. Here d is the data vector, f is the image vector, the matrix R describes the instrumental response, and σ is the estimated noise in the data. High demands are placed on the quality of the data and the response matrix R . Therefore, HIRAS utilizes the individual response functions of the focal plane detectors and rotates them according to the scan angle over the requested map (Bontekoe et al., 1991).

The spatial correlations in an image can be controlled via a new multi-channel method, pyramid images, in which virtually all spatial frequencies are represented. In a pyramid image a 64×64 pixel image is the sum of a 64×64 , a 32×32 , ..., a 2×2 , and a 1×1 pixel channel, all covering the same map area. The final result, of course, is a single map.

In image reconstruction there is an inevitable amplification of the noise in the data, due to the ill-conditioned nature of the inversion of the imaging equation. In HIRAS this amplification of the noise can be quantified by the computation of an error map, having identical dimensions to the requested map. The error map represents a full propagation of σ through the inversion of the imaging equation. Comparing the image with the error map gives indispensable information about the authenticity of detailed features. For example, the flux of point sources in the IRAS Point Source Catalog is usually within the error as determined by HIRAS.

In Groningen the complete IRAS database (survey, AOs, and LRS) has been written to a magneto-optical jukebox system, and is made accessible to the astronomical community. The extraction of IRAS data is performed by an automatic mail request server (for info: irasman@sron.rug.nl). The data are stored as integer numbers just as they were relayed from the satellite. The selected data are retrieved in the form of a IRAS Data Set (IRDS) (Roelfsema and Kester, 1992), which is a specialized version of a GIPSY data set. The Groningen Image Processing SYstem (GIPSY)³ is a general purpose astronomical image processing package (van der Hulst et al., 1992). In GIPSY there are many software tools for processing IRAS data.

The IRDS is the basis of all our high-resolution processing, starting with calibration from integer data numbers to MJy/sr and adding pointing information.

The data d , still administered in the form of scans, are first flatfielded and destriped against a low-resolution map (CoAdd) by the IMAGE task in GIPSY. After a first convergence of HIRAS subsequent refinements in the calibration can be performed, now against the new high-resolution map. Corrections in the baseline, drift, and detector gain are applied and continuation of the HIRAS run shows a significant improvement in the result.

HIRAS is a highly interactive GIPSY task, although it can also be run with reasonable settings in an automatic mode. Being part of a larger image processing system opens venues for inspecting and investigating the data and moulding them to obtain the best possible solution. Since HIRAS uses proprietary software, it cannot be distributed without proper licensing.

For a full description of HIRAS we refer to Bontekoe, Koper & Kester (1993). Below we describe our IRAS model, which includes corrections for the

³GIPSY can be obtained, free of charge (for info: kgb@astro.rug.nl).

noise due to digitization and data compression, de-glitching and de-tailing of the data and the redetermination of the 100 μm detector response functions.

PREPROCESSING

After calibration to physical units and the subtraction of a zodiacal emission model from the data, the IMAGE task performs the local flatfielding and de-stripping of the data. Generally one starts with a zero-valued template map. First order baseline corrections, zero points and drifts, are derived for each detector scan by fitting the scans to the template. The fit is a lower envelope fit. For detector response functions the nominal rectangular apertures are used as a first order approximation. The thus corrected scans are co-added into a new map, the CoAdd. In subsequent iterations the CoAdd of the previous iteration serves as a template map (Wesselius et al. 1992). Typically, the procedure requires five iterations, after which, in general, no further improvement can be obtained. The de-stripping works best when at least two sets of scans are present with an appreciable angle between them, but it does not exclude proper de-stripping for areas with almost parallel scan only. Some more care is necessary in such cases. The derived correction parameters are stored at the appropriate levels in the IRDS. HIRAS uses this improved IRDS as input, not the CoAdd.

ZERO POINT OF DATA

As a consequence of the above improvement of the internal consistency of the data, our knowledge of the value of the absolute zero point of the data has eroded. Ideally, a black sky should give zero signal apart from instrumental noise, which can give positive as well as negative data values. An erroneous off-set in the zero point shifts the balance between noise and data. MEMSYS5 yields non-negative images, but accepts negative data values. Negative data values are regarded as noise, in full agreement with Gaussian statistics. Gross errors in the zero point of the data result in areas in the image devoid of any emission in the case when significant data are shifted below zero. Positive offsets are not very apparent in the results, but generally show up as unrealistically high backgrounds.

In a trade-off the fiducial zero point is taken such that typically five percent of the data become negative. ‘Hard zeroes’ in the resulting image are usually avoided in this way.

NOISE MODELING

The noise in the data is estimated by application of a zero-sum filter over a sufficiently long part of the detector scan, i.e. data from a single detector from a part of a single scan over the image area. The median value of the filter output is appropriately scaled to the correct standard deviation in the case of Gaussian noise (Bontekoe et al. 1991). This yields a noise estimate σ per detector scan; to this noise estimate two corrections for non-Gaussian effects are added.

First, a contribution is added which accounts for the lossy data compression scheme on board the satellite. The bulk of the data were transmitted as

logarithmically compressed 8-bit differences, with every four seconds a 16-bit fiducial point (IRAS Explanatory Supplement 1988, App. II.1). These data were decompressed at the ground station into 16-bit integers. In the worst case, strong gradients were transmitted by only three significant bits plus a sign bit. To accommodate for this loss in precision we adjust σ as follows:

$$\sigma_n = \sqrt{\sigma^2 + a * |d_n - d_{n-1}|^2}. \quad (2)$$

($d_n \pm \sigma_n$ is the value of the n -th datum and its uncertainty.) The constant a is usually taken 0.03, corresponding to a precision loss of a few percent. MEMSYS5 allows each datum to have a different noise estimate. Subsequently, the values of consecutive noise estimates are smoothed a little to avoid too large variations.

Second, for samples which partly overlap the map boundary an additional correction for the noise estimate is made. Samples are included only if at least half of the ‘volume’ of the RF is inside the image area. The value of d_n is multiplied by the fraction of the volume inside the image area and the σ_n is divided by it, thus diminishing the relative influence of the ‘unknown world outside’.

And finally, the minimum step in the sample values, after calibration, is identical to the (scaled) unit step as when the samples were represented by integers. This step size is quite large, either of the same size as the noise estimates σ . This means that the smallest deviation possible in the data is of the order of 1σ . The assumption of Gaussian noise then tends to break down. This is similar to a Gaussian fit of a random series of $+1, -1, \dots$, which has a standard deviation of $\sqrt{0.5} < 1$. Every deviation is at least a ‘one- σ -detection’.

MEMSYS5, being designed for Gaussian noise, is quite sensitive to such systematic effects. A more fundamental approach, which transcends the Gaussian assumption, is not (yet) possible. We ameliorated the data by smoothing the least significant bit, from the 16-bit integer representation, with a cubic spline function.

TIME DOMAIN FILTERING

Thus far only global corrections on d are made in the preprocessing stage and instrumental effects are modeled in σ . However, glitches and memory effects in the data are short duration phenomena and have to be corrected in a few data samples. Glitches, originating from e.g., cosmic ray hits, are identified by application of a glitch filter and these samples are removed. Memory effects (also known as hysteresis), resulting in decaying offsets after passing over a strong compact source, are mainly present in $12\ \mu\text{m}$ and $25\ \mu\text{m}$ data.

A combination of filters and flags are implemented to signal their occurrence. Two zero-sum filters run over the data. One is mainly sensitive to point sources $\{-1, -1, -1, +2, +2, +2, -1, -1, -1\}$ and the other is more sensitive to glitches $\{-3, +6, -3\}$. (Glitches are of shorter duration than point sources.) If the point source filter output surpasses a given threshold and if it is also larger than the glitch filter output, then the source is flagged as a point source. Otherwise, it can be flagged as a glitch and the corresponding section of the data is subsequently ignored.

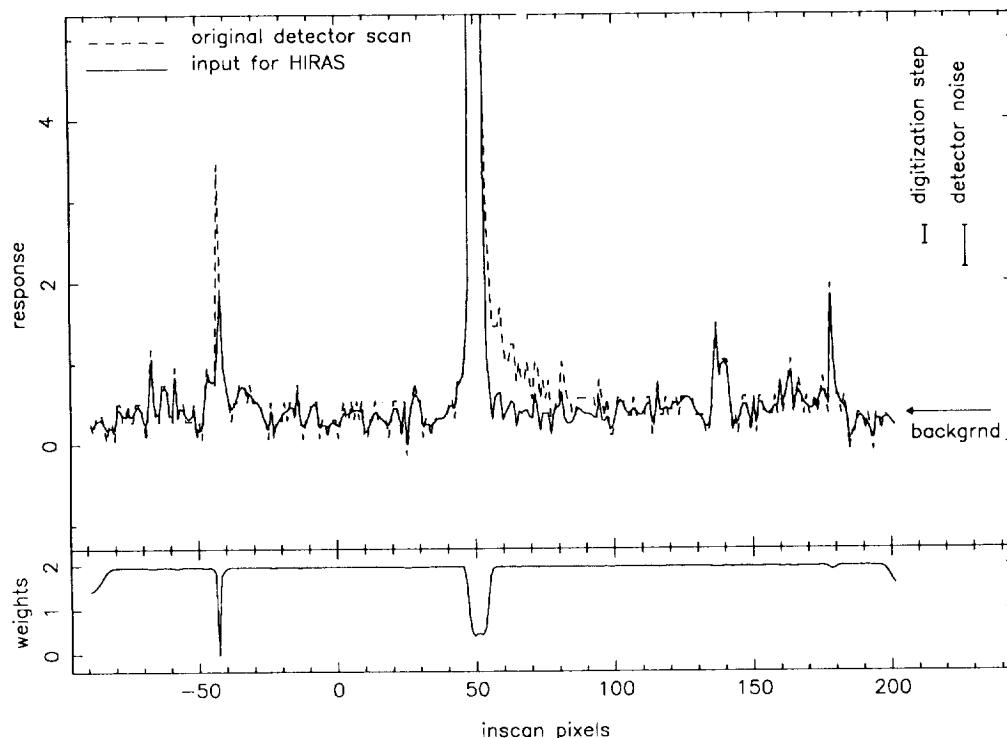


FIGURE 1 Sample values of one detector scan, running over a point source in the Chamaeleon region. In the upper panel the dashed line represents a calibrated and destriped detector scan; the full line represents the same detector scan after de-tailing, de-glitching, and digitization-smoothing, and will be input for HIRAS. In the lower panel the weights of the data are displayed, the weights are $1/\sigma_n$. A weight equal to zero indicates a sample which is rejected, e.g. a glitch.

It is well known that the IRAS detectors exhibit memory effects after having observed a bright source, especially at 12 and 25 μm . A number of data samples taken just after observing a strong point source can suffer from those nonlinearity effects, which are called ‘tails’. Depending on the strength of the filter outputs a number of tail samples are flagged. These tail samples are either excluded or corrected via a de-tail filter. We implemented a version of Russ Walker’s (priv. comm.) de-tailing algorithm, which assumes an exponential decay. We are studying a way to derive parameters of the memory effects within the MEMSYS5 context. Figure 1 summarizes the effects thus far described.

RESPONSE FUNCTIONS

The current IRAS Response Functions (RF) have been derived from a series of special raster observations on the planetary nebula NGC 6543 by Mehrdad Moshir (priv. comm.) in 1986. They were never intended for the construction of maps, though they served a good role in doing so. These observations were taken

at half survey speed. In the cross-scan direction the wings are not defined below $\approx 15\%$ of the maximum (see Fig. 2). This can be improved somewhat extending the RFs with a linear roll-off in the cross-scan direction over $\approx 1'$. This made 12, 25, and 60 μm IRAS results more acceptable when compared with, e.g., ground-based observations. Some 100 μm results, however, were inconsistent with their 60 μm counterparts. Although an exact correspondence cannot be expected, many of the physical processes are well understood; for example, the nearby edge-on galaxy NGC 55 had a significantly different morphology in the two bands. Detailed analysis of the 60 and 100 μm data led to the conclusion that the 100 μm RFs were suspect (Bontekoe, Koper & Kester, 1993).

We decided to rederive the RFs for the 100 μm detectors from survey data passing over NGC 6543. This would yield RFs at full survey speed. Since NGC 6543 is almost exactly in the pole of the IRAS orbit, it has been observed many times. From the data server 107 scans in the neighborhood could be extracted, for which all the necessary calibration and position parameters are present, and in which the source is present. In this area of the sky the on-board calibration also took place, by illuminating the focal plane with two flashes of known intensity. These flash-data were later used for the calibration of the data. Almost all scans over NGC 6543 contain these flashes, and the useful parts had to be selected manually. (Under standard operation of the data server calibration flashes are removed.) However, the focal plane is very unevenly covered by passes of NGC 6543 (see Fig. II.D.9 of *IRAS Faint Source Survey Explanatory Suppl.*). Large gaps with no useful data, in one case almost $2'$ wide, occur in the cross-scan covering density.

A response function can be derived via the following method. The imaging equation (1) has a certain symmetry in R and f . The rows of the matrix R codify how the RF of a detector, at a single instant, overlies the sky. The RF, however, can also be regarded as an image of the detector response. The datum d_n thus is the vector product of the 'image' of the detector with f . In other words, the rows in R can be regarded as shifted and rotated images of the detectors. If we postulate that we know the scene f , we can de-shift and de-rotate the image f into the rows of a new matrix \hat{R} , and solve for the image of the detector RFs. In our case, we have to decide what NGC 6543 looks like at 100 μm .

Optically, NGC 6543 is a bright II II region of approximately $10''$ in radius, surrounded by a weak halo of $190''$ radius (Millikan, 1974). From many crossings over the Low Resolution Spectrograph (LRS) it could be established that in the wavelength range from 4 to 22 μm NGC 6543 is smaller than $5''$. However most of this is line emission from the central source. Moseley (1980) attempted to determine its size in 37 μm continuum radiation, and found little flux outside a beam of $20''$. With the effective CPC beam of $88''$ and $100''$ for bands at, respectively, 50 μm and 100 μm , the nebula seems unresolved. In the CPC images there is no evidence that a halo is present (Wesselius et al. 1985). From this we conclude that the nebula must be appreciably smaller than the effective CPC beam. NGC 6543 is extended at a wavelength of 6 cm with a FWHM of $15''$ (see Pottasch 1984), which is significantly less than the IRAS telescope diffraction limit of $100''$. Therefore and also for simplicity reasons, we assume NGC 6543 to be a point source at location ($17^{\text{h}}58^{\text{m}}33^{\text{s}}.9 + 66^{\circ}38'6''$, eq1950). (Note that the IRAS PSC differs from this position by $48''$!)

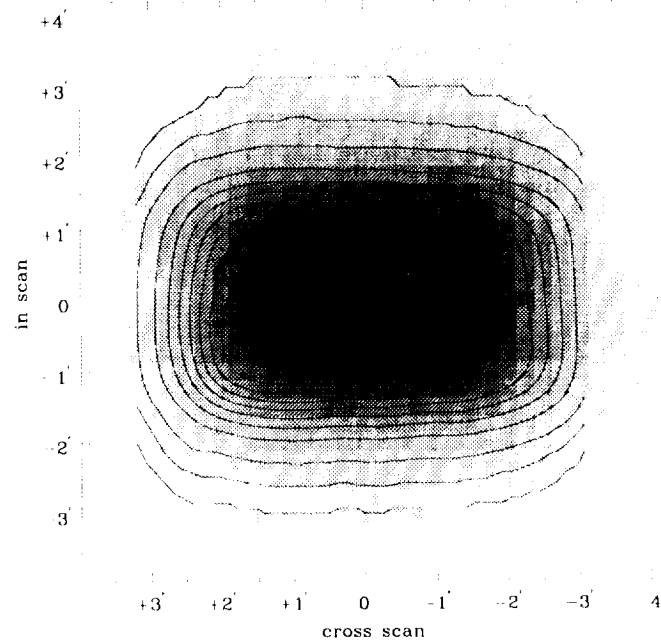


FIGURE 2 For detector 4 the new RF is displayed in graytones. It is overlaid with contours of Moshir's RF. Both contours and gray scales are at 2, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, 98% of the maximum. It can be seen that in first order the new RF is a shifted version of Moshir's RF.

Figure 2 shows the old and new RF for detector 4.

DE-STRIPING WITHIN HIRAS

After a first convergence of HIRAS, mock data \hat{d} are drawn from the image f , using the imaging equation. These data are compared with the real d , on detector scan basis. Whenever an offset or a drift could be determined with enough statistical significance, the real data d are corrected. HIRAS then continues with the improved data. This extra destriping operation gives a noticeable improvement in the results.

In addition, significant detector gain variations can be present, which are seldom seen in the data preprocessing using CoAdds. The reasons for this are twofold. First, the mock data from HIRAS, using the detector RFs, can follow the structure in the image better than when using rectangular apertures, as is done in CoAdds. Second, occasionally gain corrections arise from the steep flanks

of the RFs. Some RFs rise 2% per arcsec in the cross scan direction. The data have a pointing accuracy of 10" (*IRAS Explanatory Supplement* 1988), which combined with our standard pixel size of 15", gives a worst case error of 50% in the mock data flux. This in turn would be translated into a large gain correction factor. This de-stripping procedure is repeated until no significant deviations are present anymore.

CONCLUSION

Some results obtained with IIRAS can be seen in the contribution of Waters et al. at this workshop (pgs. 111-118).

We know that in some areas our knowledge is still incomplete, e.g. we would like to have a better model for the memory effects and new response functions for the other wavelength bands. In addition, a more fundamental solution to the digitization noise could improve things further.

IRAS performed much better than expected. We are still discovering new systematic and thus modelable effects. We are not yet down to the random noise of the instrument.

Currently we are planning to make the IIRAS program accessible via the mail server, which already is in place for coadded images and LRS spectra. More information can be obtained from irasman@sron.rug.nl.

ACKNOWLEDGMENTS

Many people have been involved in the development of the IRAS software. We would like to thank Timo Prusti, Pjotr Roelfsema, Albrecht de Jonge, Rens Waters, Rob Assendorp, Freddy Lahuis, Enrico Koper, Paul Wesselius.

REFERENCES

- Bontekoe, Tj.R., Kester, D.J.M., Price, S.D., de Jonge, A.R.W., and Wesselius, P.R. 1991, *A&A*, **248**, 328
- Bontekoe, Tj.R., Koper, E., and Kester, D.J.M. 1993, in preparation
- van der Hulst, J.M., Terlouw, J.P., Begeman, K.G., Zwitter, W., Roelfsema, P.R. 1992, in *Astronomical Data Analysis and Software Systems I*, D.M. Worrall, Ch. Biemesderfer and J. Barnes (eds.), Astronomical Society of the Pacific Conference Series, **25**, 131
- IRAS Catalogs and Atlases: Explanatory Supplement* 1988, ed. C.A. Beichman, G. Neugebauer, H.J. Habing, P.E. Clegg, and T.J. Chester (Washington DC: GPO)
- Millikan, A.G. 1974, *AJ*, **79**, 1279
- Moseley, H. 1980, *ApJ*, **238**, 892
- Moshir M., et. al. 1992, *Explanatory Supplement to the IRAS Faint Source Survey*, (Pasadena: JPL)

- Pottasch, S.R. 1984, *Planetary Nebulae*, Reidel, Dordrecht
- Roelfsema, P.R. and Kester, D.J.M. 1992, in *Astronomical Data Analysis and Software Systems I*, D.M. Worrall, Ch. Biemesderfer and J. Barnes (eds.), Astronomical Society of the Pacific Conference Series, **25**, 359
- Wesselius, P.R., Beintema, D.A., de Jonge, A.R.W., Jurriens, T.A., Kester, D.J.M., van Weerden, J.E., de Vries, J. and Perault, M. 1985, *Chopped Photometric Channel Explanatory Supplement*, SRON-SRG, Groningen
- Wesselius, P.R., de Jonge, A.R.W., Kester, D.J.M., and Roelfsema, P.R. 1992, in: *Infrared Astronomy with ISO*. Encrenaz T., Kessler M.F. (eds.) Nova Science Publishers, New York, p. 509

